

**Good and Bad News About Florida Student Achievement:
Performance Trends on Multiple Indicators
Since Passage of the A+ Legislation
Policy Brief**

Madhabi Chatterji

Columbia University

Education Policy Research Unit (EPRU)
Education Policy Studies Laboratory
College of Education
Division of Educational Leadership and Policy Studies
Box 872411
Arizona State University
Tempe, AZ 85287-2411

April 2004

EPSL | EDUCATION POLICY STUDIES LABORATORY
Education Policy Research Unit

EPSL-0401-105-EPRU
<http://edpolicylab.org>

Education Policy Studies Laboratory
Division of Educational Leadership and Policy Studies
College of Education, Arizona State University
P.O. Box 872411, Tempe, AZ 85287-2411
Telephone: (480) 965-1886
Fax: (480) 965-0303
E-mail: epsl@asu.edu
<http://edpolicylab.org>

**Good and Bad News About Florida Student Achievement:
Performance Trends on Multiple Indicators
Since Passage of the A+ Legislation**

**Madhabi Chatterji
Columbia University**

**Good and Bad News About Florida Student Achievement:
Performance Trends on Multiple Indicators
Since Passage of the A+ Legislation**

Madhabi Chatterji
Columbia University

Executive Summary

What is the news on Florida's student performance since the passage of the A+ legislation? This brief presents a review of long-term data on five state and national indicators. It verifies outcomes and trends and examines the main premise of the A+ mandate that, given appropriate schooling, students will have equitable outcomes and access to opportunities. The data show positive outcomes and steady gains at the elementary level; the pattern, however, is not sustained at the secondary level.

Good News at Elementary Level in all Subjects

Elementary achievement trends show the strongest gains and performance in writing, and steady improvements over time in reading and mathematics. Findings concur on the FCAT-CRT, FCAT-NRT, and the NAEP.

Bad News on Secondary Level Reading

About 60% of Florida's high schoolers are still performing below the Level 3 (passing) mark on the FCAT-CRT in 10th grade, while less than 50% place below the national median on the FCAT-NRT.

Good News on Writing and Fair News on Mathematics at all Levels

On the FCAT writing tests, more than 80% stayed above Level 3 for three to five years in all grades. Compared to other U.S. students in all grades tested over time,

Florida's students have an advantage in writing, with 84-86% at or above the Basic level in the most recent NAEP testing. At all levels, Florida's students also steadily improved over time in mathematics achievement on the FCAT-CRT, FCAT-NRT, and the NAEP.

Mixed news on Achievement Gaps among Florida's Students

Achievement gaps are still large between White and other subgroups as of 2003, but appear to be smaller in some subgroup comparisons, such as White vs. Hispanic students (see 2003 results in [Appendix D](#)). On the FCAT-CRT in 2003, Limited English Proficient students and Exceptional Education students showed the poorest performance in all grade levels. Also on the FCAT, among elementary school students, ethnic achievement gaps that were slowly closing over time in reading showed no change. The data revealed widening gaps in secondary grades: a large female-to-male gap in reading starts in elementary school and widens in middle school on the NAEP. A small male-to-female gap in mathematics is evident in elementary school, and also widens in middle school (NAEP). On the NAEP, among Florida's students, gaps between ethnic minority groups and Whites remained the same or decreased in mathematics over time; the gaps remained the same or increased on reading and writing tests over time.

Mixed News on Graduation Rates, Dropout Rates, and College-bound Seniors

Over time, U.S.-to-Florida comparisons show a gap of 10-12% in the high school graduation rate per year through 2000: Florida is behind, with a graduation rate of 55% in 2000. The *number* of high school diplomas issued in Florida, however, has increased over time. Florida's dropout rate, based on counts of students giving withdrawal reasons, has slowly declined from 5% to 3% since the A+ legislation. Calculations of both graduation and dropout rates are limited as they do not include student migration or grade

retention factors. More college-bound seniors from Florida are now taking the SAT, with a greater percentage from ethnic minority groups. U.S.-to-Florida comparisons show a widening gap in both quantitative and verbal areas. Only 25% of Florida's test-takers reported that they were from high-income families in 2003, compared to 58% in the nation. Based on the documented relationship between income and SAT performance, lower scores are predictable when students come from low-income households. On the SAT composite in 2003, however, Florida's Hispanics scored +39 points better on average than Hispanics in the nation. More information on these data may provide insight as to what factors affect minority student success.

Given the current data trends and research, it is recommended that Florida legislators and policy-makers undertake the following:

1. Provide local support and evaluation training opportunities to help teachers and school leaders make better use of data to reduce achievement gaps in reading and mathematics, particularly at the secondary level.
2. Document and reward the pedagogically sound practices that follow from Recommendation 1, even when short-term gains are small. These practices, informed by data, will support student learning and sustained growth.
3. Create conditions to help schools teach to the Sunshine State Standards along a K-12 continuum – conditions that lead to real engagement of teachers and students in learning activities tied to standards and that reduce the amount of test preparation. This includes removing the use of teaching to the (FCAT) test language, currently on the A+ website, that tends to encourage short-term gains.

4. Make greater use of appropriate sampling techniques (NAEP type of design) in school evaluations, lowering anxiety and fear resulting from the present accountability system. This will allow Florida to retain high levels of accountability while lowering the stakes for individual school staff.
5. Seek ways to set reasonable and staggered timelines, within the NCLB and A+ frameworks, for schools with diverse and challenging populations to demonstrate adequate yearly progress (AYP).
6. Rethink sanctions and grade retention policies in light of new and past research that finds that grade retention does not improve student achievement.
7. Identify specific reforms with positive and sustained achievement effects by supporting long-term research that looks at diverse students' achievement in relation to school and teacher practices over time.

Good and Bad News About Florida Student Achievement:

Performance Trends on Multiple Indicators

Since Passage of the A+ Legislation

Madhabi Chatterji
Columbia University

Section 1: The Issue

Since the mid-1980s, the standards-based education reform and accountability movement has significantly raised expectations of performance for U.S. schools.¹ In 1999, in order to improve K-12 education in Florida, Governor Jeb Bush promoted the A+ legislation, a standards-based accountability system that assigns grades of A-F to schools. The A+ law requires that students in grades 4-5, 8, and 10 annually take criterion-referenced assessments in reading, mathematics, and writing, as a part of the *Florida Comprehensive Assessment Test* (FCAT) program.² The FCAT now also incorporates a norm-referenced test in grades 3-10 in reading and mathematics.

As the five-year mark approaches since the passage of the A+ legislation, what is the news on Florida's student performance? Do student achievement trends suggest that Florida's standards-based reform strategies have worked? Do trends concur on the various available indicators of student performance, such as state tests (FCAT), national tests (*National Assessment of Educational Progress*, NAEP), high school graduation rates, and tests for college-bound seniors (*Scholastic Assessment Test*, SAT)? The Florida law aimed to reverse low achievement trends and reduce existing achievement gaps. Therefore, what is the present status of the gap between Florida and the nation, and among various student subgroups within Florida?

To answer the preceding questions, this policy brief examines trends according to five selected indicators of student performance. The evidence is derived from published state and national reports and has been evaluated against research findings, commentaries, and recent federal legislation on high-stakes, standards-based educational reforms and accountability. Based on the evidence, the brief concludes with a series of recommendations for improving Florida's schools and student achievement levels.

Section 2: Background

Legislative Trends

Standards-based Reforms and the A+ Accountability Law in Florida

The 1994 re-authorization of the federal Elementary and Secondary Education Act mobilized resources for standards-based reforms on a national scale, channeling money to individual states.³ Around that time, educators and the public in Florida developed and adopted a rigorous set of subject area standards, called the Sunshine State Standards (SSS). The SSS, developed under Governor Lawton Chiles' leadership, specified what students should know and be able to do in grades K-12, and served as the framework for the development of the new *Florida Comprehensive Assessment Test* (FCAT) program in reading and mathematics. Since 1993, *Florida Writes!*, a timed essay-writing test, has been administered annually in grades 4, 8 and 10; by 1998 it was incorporated within the new FCAT program. In 1998, the State Board of Education approved five achievement levels for placing students in different performance categories based on FCAT scores. All three tests in the original FCAT program were thus "criterion-referenced," permitting interpretations of student performance by comparison

with a set standard or “cut score” on the test, rather than by ranking students among a group of their peers.⁴

Shortly after, in the fall of 1998, the Bush/Brogan governorship assumed office and obtained legislative approval for A+, dubbing it “a comprehensive system of school reform.” The main principles of the A+ plan were that each student should “gain a year’s worth of knowledge in a year’s time in a Florida public school” and that “no child will be left behind.” Hence, the A+ law’s major focus is improving performance of *all* low-achieving students, irrespective of their ethnicity, gender, poverty levels, native language, or disability status. The state’s role is to measure results and hold schools accountable to law and current policy. Various rewards and sanctions are attached to the FCAT results and school grades. In high school, passing the FCAT with scores at a pre-set level is a graduation requirement. Because of the serious consequences tied to state test scores and school grades—consequences affecting both schools and individual students—Florida’s school assessment and accountability program is a “high-stakes” one.⁵ Other than a reduction in class size as stipulated by a 2002 amendment, the key components of the A+ accountability plan remain as they were in 1999, and include the following:⁶

1. **Equitable student gains on the FCAT:** Student achievement is the centerpiece of the A+ plan. It is monitored annually by means of the FCAT, and disaggregated by group, on both the original criterion-referenced tests (FCAT-CRT) and a new “norm-referenced” test (Stanford 9, FCAT-NRT), which offers a scale score for longitudinal tracking and allows comparison of student achievement with a national norm group. Because the originally developed FCAT tests were criterion-referenced, the new plan called for a

transition to the new FCAT-NRT that would eventually enable mapping of student achievement in grades 3-10 using a “value-added” approach.⁷

2. **School Grades:** Schools are awarded grades A-F based primarily on students’ FCAT scores. The FCAT scores of schools with similar mobility and poverty rates are compared, and other indicators (out-of-school suspension, absenteeism, dropout, and promotion) also are considered. As of now, not enough information is available from the state regarding the exact procedures and weights allocated to different factors in calculating school grades.
3. **Additional Funding/Rewards/Sanctions:** Low-performing schools are publicly identified and initially provided with financial support for a variety of programs of their choice, such as teacher recruitment and retention, mentoring, tutoring, or class size reduction (before 2002), to raise student achievement levels. Schools showing gains of one letter grade are recognized through additional funding; state law also requires that teachers be awarded merit pay based on performance.
4. **School Vouchers:** If after two years of additional resources from the state, schools fail to improve, parents are provided with “opportunity scholarships” so they may choose other private or public schools for their students.
5. **Student Retention:** The law states that students not meeting specific levels of performance should receive remediation or “be retained within an intensive program” that differs from the previous year’s educational program, or both.

The No Child Left Behind Act of 2001(NCLB) followed the passage of Florida’s A+ legislation by a couple of years.⁸ The parallels between the two laws are

unmistakable, and a 2001 *New York Times* article noted that Governor Bush's plan was held up nationally as a model for combining aggressive testing with a voucher program to bring about significant education reform.⁹ A brief history of the national law and of existing research on key strategies in the reform plan follows.

Standards-based Reforms and Accountability Legislation in the Nation

In the U.S., sweeping education reforms were sparked by the April 1983 report of the National Commission on Excellence in Education, which stated that the nation's youth were not learning enough and education systems in the U.S. needed immediate improvement.¹⁰ Those reforms continue today with the recent passage of the federal No Child Left Behind Act of 2001.¹¹ Like the A + plan and earlier legal measures, such as the 1994 re-authorization of the Elementary and Secondary Education Act, state reforms prompted by NCLB include four main directives: rigorous *content standards* set by states (difficult subject matter and higher order thinking skills); *standards-based tests* to measure and document student and school performance on the new content standards; *performance standards* (bars that separate students into performance categories based on the new standards-based tests); and *rewards/sanctions* for schools and students, based on how well they meet the performance standards set by individual states.¹²

The NCLB Act, approved by a bipartisan Congress, is sweeping in its aims and provisions. Among a vast array of issues, the Act addresses the teaching of reading, family literacy, delinquency and dropout prevention, teacher training, language instruction for immigrant students, charter schools, and the adoption by schools of practices that are supported by scientifically based research.

The Act's main strategy for school improvement, however, consists of implementing higher academic standards coupled with high-stakes testing and

accountability. For the first time in the history of national reforms, the law emphasizes achievement by *all* groups of students, particularly those who are historically low-achieving, such as socio-economically disadvantaged or special needs students. As in the A+ plan in Florida, testing requirements have substantially increased to include all students in grades 3-8, plus students in an additional year in high school. Results of student performance are to be broken down by relevant subgroups, and states are required to follow a precise time-line to close the achievement gap between different groups by 2014. The Act includes a uniform “one-size-fits-all” mandate for grade levels and subjects to be tested; procedures on setting starting points for longitudinal tracking of schools and students; and procedures for setting objectives in order to make Adequate Yearly Progress (AYP) towards the 2014 goal. Consequences for poorly performing schools vary in severity and will be enforced over periods of two to five years. These include provision of technical assistance from the state; options for transfers and public school choice for students; funding for supplemental instructional services; shifts in staffing; and school takeovers or governance changes.¹³

High Stakes Reforms and Accountability: Reactions from Educators and the Public

Good and Bad Aspects of High-stakes Reforms

Many educators and parents support raising standards. The American Federation of Teachers (AFT) has endorsed the move towards higher standards in subject areas, and AFT leaders hold that establishing clear and consistent standards across the nation is likely to “help alleviate the frustrations associated with student mobility.”¹⁴ National teacher and parent surveys conducted in 2000-01 also show that a vast majority of public school parents and teachers agree that raising academic standards is “a move in the right

direction.” Some reports say that despite pressures, teachers and the public still believe that standards, assessment, and accountability will improve schools.¹⁵

By 2002, nearly all fifty states adopted new content standards in two subjects: reading/language arts and mathematics. Content standards in particular subjects were so numerous in some states, however, that a call to streamline them ensued.¹⁶ The clarity, specificity, and difficulty levels of content standards also varied greatly across different states.¹⁷

States also began to set performance standards, several following the NAEP approach, which recognizes four levels of proficiency. Like the content standards, minimum performance levels varied from state to state in terms of the range of targeted knowledge and skills, the stringency of standards, and their comparability with minimum levels set on other national tests such as the NAEP.¹⁸ Differences were evident even within the same subject areas.

The most problematic aspect of the NCLB legislation, however, concerns the expectations for individual school progress as well as for closing the individual achievement gap between various student groups by a pre-set deadline. Some educational researchers argue that the requirements are unrealistic. As students begin to take the new standards-based tests, schools in different states are certain to vary considerably in where they start. Statistical projections suggest that, to meet the 10-year goals for improvement required under the act, schools would have to produce impossibly steep rates of achievement growth.¹⁹ Further, the projections suggest that schools that do perform well (i.e., show small, steady gains, but fail to fully close the achievement gaps) may not be recognized by the NCLB system of accountability.²⁰ One cannot realistically assume that the student demographics and other contextual variables in schools, districts,

and states will remain constant during the required time period for meeting the NCLB goal. Thus, the chances are high that hard-working and effective schools will face harsh sanctions despite making small but definitive gains.

Research Evidence on Reform Strategies in the A+ and NCLB Laws

Effects of High-stakes Testing on Classroom and School Practices

The research evidence shows clearly and repeatedly that even when teachers believe in higher standards, they respond to high stakes accountability systems by narrowing their classroom curriculum, particularly when under external pressure. The subjects covered on high-stakes external tests are emphasized during instruction, while non-tested subjects are omitted. Another negative consequence is the dedication of excessive teaching time towards test preparation activities—that is, drilling and practice on particular test items or particular item-formats—as opposed to teaching the broader knowledge and skills in the curriculum.²¹

In a recent national study of teachers in 47 states, including 167 elementary and secondary teachers from Florida, 70-80% of teachers across the states reported that they prepared students for the state-mandated test throughout the year, and 62% of the Florida teachers agreed that they greatly increased instructional time on areas tested on the FCAT. Although this was not a scientific poll, the findings concur with other reports on the same topic. The same study showed that 90% of the Florida teachers agreed that the state testing program led teachers to teach in ways that contradict their own ideas about sound practice.²² When such “teaching to the test” practices predominate in schools and classrooms, gains on the accountability tests may be spurious and unstable, and should be validated against other indicators of student and school performance.

Under pressure to demonstrate student achievement, principals and school leaders respond in similar ways. When interviewed, 50 Florida school principals and 25 central office personnel were found to engage in the following activities soon after the passage of the A+ law: setting urgent goals to improve test scores; aligning instruction to the state standards and tests; emphasizing academic subjects that were tested over those that were not; strengthening professional development primarily on state tests and standards; and examining student achievement results on state tests.²³

Some say that the high-stakes testing and accountability phenomenon has led to benefits, however.²⁴ According to Cizek, professional development programs for school staff are now more “results-oriented”²⁵; more attention is now diverted towards designing and accommodating tests for special populations; teachers and schools have learned more about testing and assessment practices; accountability systems and tests have improved because of greater public scrutiny; and students are probably learning more as a result of the altered standards and tests.²⁶

Evidence on Reform Implementation Practices: What Schools Do Matters

The reforms that school districts and schools implement, even in a high-stakes environment, have an impact on the outcomes they achieve. When activities focus too narrowly or singularly on external tests, undesirable and often unstable outcomes can be expected. On the other hand, when time, people, materials, and training are allocated sensibly towards reforms, better continuing outcomes are reported. Effective practices and policies include aligning the local curriculum and instruction to state content standards along a K-12 continuum, with resources/supports for both classroom teachers and students²⁷; adapting standards-based instruction to meet the developmental needs of diverse students²⁸; allocating reasonable time for teachers and schools to implement

reforms²⁹; providing greater community-building and professional development around the broader content standards/assessment practices³⁰; employing regularly licensed and more experienced teachers to teach students, rather than those who are emergency-certified³¹; making improvements based on continuous monitoring of student, teacher, and school progress³²; and using standards, rather than the tests, as the main catalyst for reforms.³³

Particularly, schools that succeed in reducing achievement gaps in the nation tend to make better use of data, which results in specific changes in their day-to-day practices.³⁴ Appraisal of data in schools needs to be systemic and comprehensive. To make better local decisions, rather than simply examining test scores (outcome data), schools should review data on context variables (school/student backgrounds and needs), inputs (allocation of particular resources to help students learn), and school/classroom processes (what teachers and schools actually do to improve learning outcomes), along with data on student test scores. By formally gathering and reflecting on evidence regarding which policies and practices help students achieve, schools will be in a better position to shape students' growth strategically.³⁵

Grade Retention as a Strategy for Improving Student Achievement

The student retention requirement in the Florida A+ plan is controversial. Do grade retention practices improve educational standards and future student achievement? Research on grade retention has not shown long-term achievement benefits, particularly for older students, and demonstrates that even higher attrition rates from high school may result. Such evidence has led researchers to conclude that high-stakes testing coupled with grade retention forces large numbers of low-achieving students to attend summer school or repeat a grade, increasing social and ethnic inequalities, and placing youth at

greater risk of dropping out.³⁶ When the focus is on retention, attention is diverted from determining *why* some students are failing in schools while others are not. A recent study on the effects of Florida's previous retention policies, a part of the 1980s basic skills education program in Miami-Dade County, concluded that there was little or no demonstrable improvement in achievement by retention of students at risk of academic failure.³⁷ In 1990, the Florida legislature discontinued those retention policies.³⁸ Retention is, moreover, an expensive practice. The relative cost of each additional year of schooling for each retaineer, when calculated against gains made on achievement tests, suggests that retention practices in the state may not be cost effective.³⁹

Research on National Achievement Trends and Gaps: Have Outcomes Been Equitable?

Both the A+ plan and NCLB seek equitable outcomes for *all* students. The premise is that similar achievement outcomes in students will lead to more equitable access to future education, jobs, and quality of life. Prior to examining current achievement trends and gaps in Florida, it is useful to examine the findings of some national trend studies as a framework for comparison. In national samples using two of the indicators (NAEP and SAT) selected for the present brief, gaps between ethnic groups show similar patterns over time.

NAEP Trends and Gaps

The long-term aggregate trend on the NAEP test shows that student reading and mathematics achievement improved moderately in the U.S. between the early 1970s and 1999.⁴⁰ Despite the overall improvement in that period, substantial differences existed among different racial groups in both reading and mathematics, and across all grades tested.⁴¹

The Black-White gap over time was found to reduce by 20-40% overall during the 1970s to 1999. The narrowing occurred mainly between 1971-1988, however, when Black students improved in all age groups but White students showed flat performance (differences range from about 18-30 points in 1988). Yet between 1989 and 1999, Black students dropped in performance while White students gained, the gap again becoming wider (differences range from about 27-35 points in 1999). The trend in the Hispanic-White gap on the NAEP was quite similar in the late 1980s, although the narrowing was not as great as in the Black-White gap in the late 1980s.⁴²

SAT Trends and Gaps

If one looks at the national gap-trend lines for the same years on the SAT verbal and mathematics tests, the patterns look very similar to those on the NAEP. The changes in the composition of SAT test-takers should be taken into account in interpreting the SAT evidence, but remarkably, both the Black-White and Hispanic-White gaps narrowed significantly in the late 1980s. The former narrowed more than the latter, from a little over 70 down to 55 points. The gaps in both cases widened again, however, in the next 10 years from a little over 70 up to 100+ points in 2000.⁴³

Possible Causes

What school conditions in the past caused the gaps to narrow and then to widen on the national tests? The shifts in the gap have been attributed to the shifts from basic skills instruction through the 1980s to the onset of standards-based reforms and an instructional emphasis on more challenging material in the 1990s. Factors such as increased school expenditures, increased resources for compensatory education programs, or high school course-taking patterns have not been found to account conclusively and fully for changing gaps in achievement.⁴⁴

Section 3: Achievement Trends and Gaps in Florida: Data and Findings

In this section, available data on achievement trends and the achievement gap in Florida's students are now compiled from the latest available public records, starting with in-state and followed by national tests and indicators. Other than graduation rates and performance of college-bound seniors, data from common grade levels (4-5, 8, and 10) and subjects (reading, writing, mathematics) are shown to facilitate comparison across state and national tests. The *Notes and References* section of this brief provides more information on the data sources. The indicators in this report appear in order of presentation:

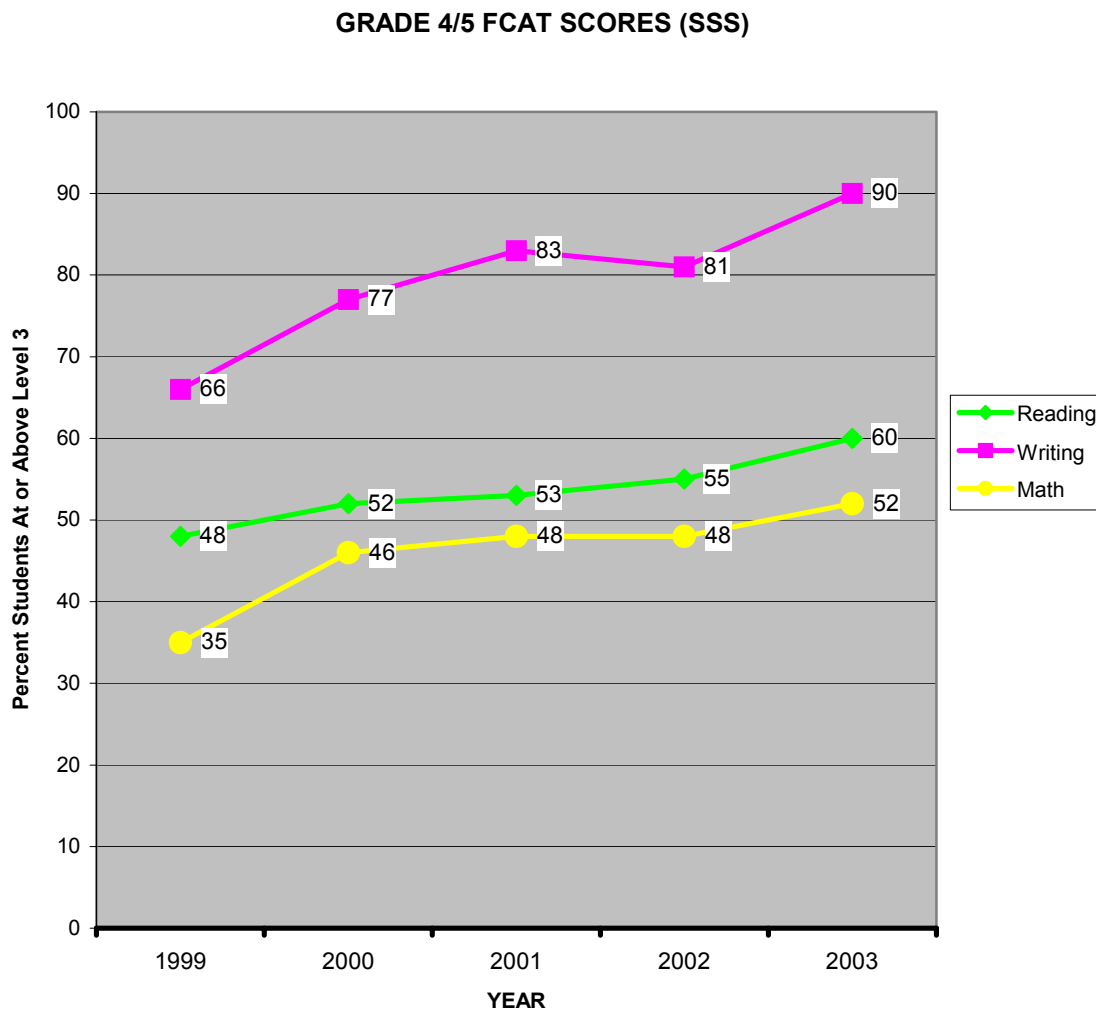
1. FCAT (SSS) CRT performance 1999-2003 (Figures 1-3, Table 1, details in [Appendices A-D](#))
2. FCAT NRT performance 2000-03 ([Appendix E](#))
3. FCAT Trends in Numbers of Test-takers (Table 2, [Appendices A-E](#))
4. NAEP Test performance 1992-2003 (Tables 3-4, [Appendix F](#))
5. Nation vs. Florida high school graduation rates 1992-2000 (Table 5) and Florida dropout rates 1999-2003 (Table 6)
6. SAT performance in Florida's college-bound seniors 1992-2003 (Figure 4, Table 7, [Appendices G-H](#)).

Performance Trends on the FCAT-CRT(SSS) 1999-2003

Elementary schoolers in Florida made strong and steady increases in Reading, Mathematics, and Writing achievement between 1999-2003, with less than 50% below Level 3 in 2003 in all subjects (Figure 1, [Appendices A-C](#)). At the elementary level, there has been a visible increase in the percentage of students scoring at or above Level 3

on the FCAT (SSS) tests in reading (from 48% to 60%) and mathematics (from 35% to 52%). In writing, particularly, above 80% of students have remained at or above Level 3 for *three* consecutive years, hitting the 90% mark in 2003.

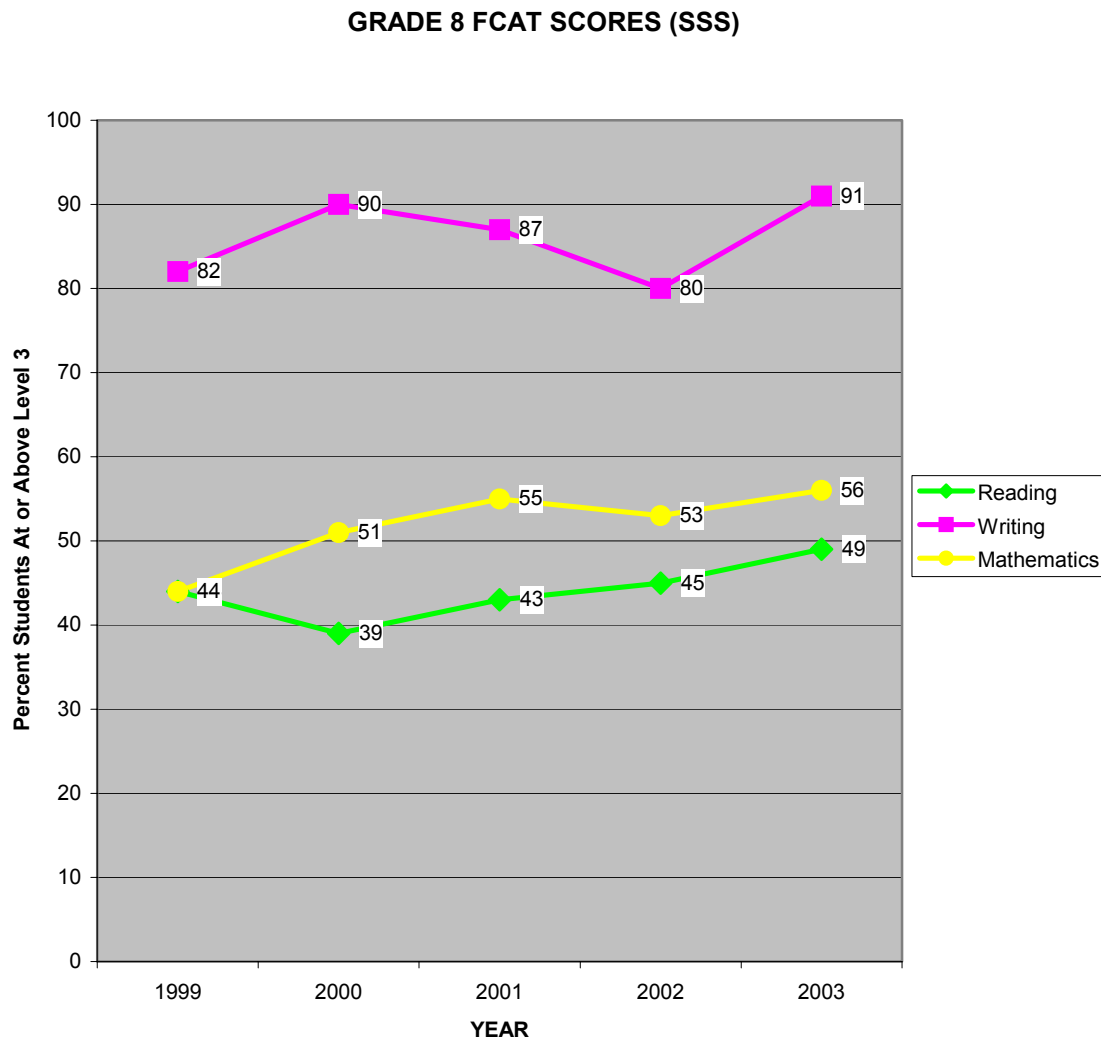
Figure 1: Elementary Students' Performance Trends in Reading, Writing, and Mathematics



Middle schoolers in Florida made small increases in Reading, Mathematics, and Writing Achievement between 1999-2003, with 51% below Level 3 in Reading in 2003 (Figure 2, [Appendices A-C](#)). At the middle school level, increases in the percentage of students scoring at or above Level 3 were not as great as those at the elementary level,

and the trend line is a little inconsistent in reading, but there is a generally increasing trend in both reading (from 44% to 49%) and mathematics (from 44% to 56%). In writing, again, more than 80% of students remained at or above Level 3 for *five* consecutive years. In 2003, 91% reached Level 3.

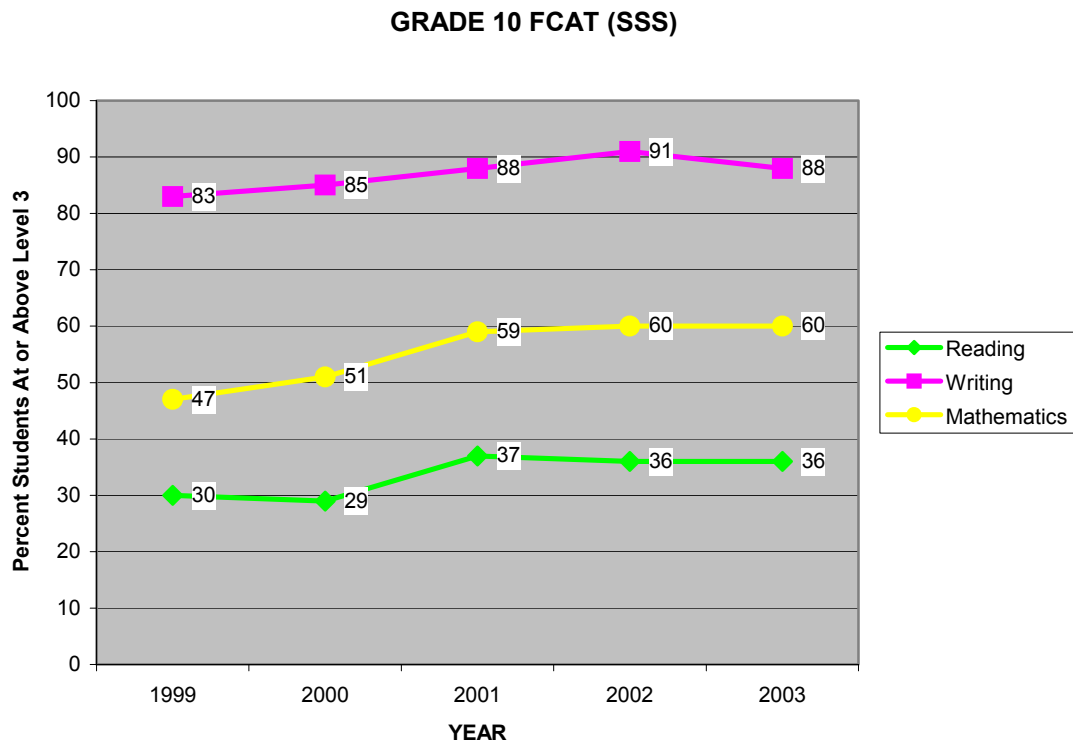
Figure 2: Grade 8 Students' Performance Trends in Reading, Writing, and Mathematics



High schoolers made steady increases in Mathematics and maintained high performance in Writing, but not in Reading. In 2003, 64% fell below Level 3 in Reading (Figure 3, [Appendices A-C](#)). At the high school level, increases in the percentage of

students scoring at or above Level 3 were the smallest compared with increases in other levels (reaching maximums of 6-7 percentage points), and, in reading, never went above the 37% mark. The reading trend line shows that for five years, 63-71% of 10th graders *were below* the passing achievement level of 3 on the FCAT-CRT. In mathematics, there is a generally increasing trend, from 47% to 60%. In writing, once more, more than 80% of high school students were at or above Level 3 for *five* consecutive years, reaching 88% in 2003.

Figure 3: Grade 10 Students' Performance Trends in Reading, Writing, and Mathematics



FCAT-CRT (SSS) Achievement Gaps

Ethnic achievement gaps were slowly declining in Reading among elementary schoolers, but gaps in other grades and subjects either did not change or else grew wider (Table 1). The Black-White and the Hispanic-White gaps seem to be declining slightly in

reading from 4-6 percentage points since 2000 through 2003, using the Level 3 cut-score criterion, but show no change or increasing trends in grades 8 and 10. A small *decline* of 4% is also evident in the grade 10 Hispanic–White Gap in mathematics. The changes, although positive, are still small and require further consideration.

Subgroup breakdowns by grade in 2003 on Reading and Mathematics (FCAT-CRT) show wide gaps with strongest performance in White students and weakest performance in limited English proficient and exceptional subgroups ([Appendix D](#)).

Comparatively, performance was highest in White, Hispanic, and low-income subgroups in all grades, using percents at or above Level 3 as the criterion. Ethnic gaps were similar in size across grade levels.

Table 1: Racial /Ethnic Achievement Gaps – Percentages of Students Scoring 3 and Above over Time

FCAT Reading					
	2000	2001	2002	2003	Direction
Grade 4					
Black – White Gap	38%	35%	31%	32%	Declining
Hispanic – White Gap	26%	23%	21%	22%	Declining
Grade 8					
Black – White Gap	35%	35%	34%	35%	No Change
Hispanic – White Gap	22%	25%	23%	24%	No Change
Grade 10					
Black – White Gap	25%	34%	33%	32%	Increasing
Hispanic – White Gap	18%	24%	23%	23%	Increasing
FCAT Mathematics					
	2000	2001	2002	2003	Direction
Grade 4					
Black – White Gap	34%	34%	33%	33%	No Change
Hispanic – White Gap	19%	19%	17%	18%	No Change
Grade 8					
Black – White Gap	39%	38%	39%	39%	No Change
Hispanic – White Gap	24%	24%	25%	23%	No Change
Grade 10					
Black – White Gap	43%	40%	41%	42%	No Change
Hispanic – White Gap	27%	24%	25%	23%	Declining

Data Source: Walberg, H. (2003) Policy Brief 3: Closing the Achievement Gap. Pre-publication Version. Jacksonville, Florida: UNF/FSU

FCAT-NRT Performance Trends 2000-03

*In Reading, on average and over time, Grade 4 and 8 students performed consistently above 50% of similar students in the nation ([Appendix E](#)). From 2000-03, the four years since testing began with the FCAT-NRTs, Florida students in Grades 4 and 8 tended to have averages (medians) that placed them *above* 54-60% of similar students in the national comparison group (see Median NPR column in [Appendix E](#)).*

In Reading, on average and over time, Grade 10 students tended to perform below 50% of similar students in the Nation ([Appendix E](#)). The bad news in Grade 10 reading performance is that the median score of Florida's 10th graders placed them at the 50th percentile in 2002, but in other years their rank ranged from the 33rd percentile to the 49th percentile, compared with the national norm group. In 2003, the Median NPR for 10th graders was 46, showing that their average score placed them above 46% of similar test-takers in the nation.

In Mathematics (2000-03), 5th, 8th, and 10th graders, on average and over time, performed consistently better than 50% -60% of similar students in the nation. The good news appears to be in mathematics performance, and in all levels on the FCAT-NRT: Florida students' average scores placed them between the 54th and 67th percentiles nationally. In 2003, the median NPR was 63 in Grade 5 (the average score placed Florida's 5th graders above 63% of similar students in the nation), 65 in Grade 8, and 66 in Grade 10. There were generally consistent increases in all three grade levels over time.

Trends in Numbers of FCAT Test-takers by Grade

In given years, numbers of FCAT CRT and NRT test-takers increased over time, and were consistently higher in Elementary and Middle School than in High School (Table 2, [Appendices A-E](#)). The pattern in numbers of test-takers in every grade level and subject area is similar on FCAT (SSS) and FCAT NRT over time. Because 95% of students are required by law to take the tests, this trend would appear to be a result of increasing public school enrollments since 1999.

In any given year, however, the numbers of test takers are markedly lower in high school than in elementary and middle school on both the FCAT (SSS) and FCAT NRT.

This last trend, denoting a drop by approximately 30,000-40,000 students as they move from elementary to high school, may be caused by student migration (caused by the A+ voucher provision or family mobility factors), or by grade retention policies or dropouts prior to high school, and thus needs further investigation. For example, comparison of the number of test-takers in reading varied in the three grade levels, as shown in Table 2 for 2001-03.

Table 2: FCAT (SSS) Reading Test-taker Numbers in Elementary, Middle, and High School by Year

	2001	2002	2003
Grade 4	188,696	191,886	193,391
Grade 8	174,016	184,483	192,116
Grade 10	144,471	150,131	167,396

Florida Students' Trends on NAEP, 1992-2003

Over time, rising percentages of Florida's elementary schoolers performed at or above the Basic Level on the NAEP Reading (63% in 2003) and Mathematics tests (76% in 2003)([Appendix F](#)). In mathematics, the NAEP tests were administered three times since 1992, while in reading they were administered five times, with greater numbers of students taking the test in later years. While only 52% of 4th graders performed at or above the basic level in mathematics and only 53% were at or above the same level in reading in 1992, the corresponding numbers were 76% and 63% in 2003 ([Appendix F](#)), showing clear gains for Florida's elementary students in both subjects, over time.

Over time, rising percentages of middle schoolers performed at or above the Basic Level on the NAEP Reading (68% in 2003) Mathematics tests (62% in 2003)([Appendix F](#)). In 8th grade, the NAEP reading and mathematics tests were

administered three times since 1992, with similar numbers of students taking the test both times. While only 49% performed at or above the basic level in mathematics and 67% were at or above the same level in reading in 1992, the corresponding numbers were 62% and 68% in 2003 ([Appendix D](#)). Here, Florida 8th grade students made an average 13 percentage point gain in mathematics, and maintained about the same level of performance in reading.

Compared to the nation and in both grades tested over time, strongest NAEP performance of Florida's students was in Writing, where they performed better than U.S. students ([Appendix F](#), Table 3). Over time, and in both 4th and 8th grades, 78-86% Florida students performed at or above the Basic level in writing. Compared with other students across the nation, Florida's students have posted higher scale scores in the area of writing. ("Scale score" is a type of score reported by the NAEP that permits calculation and comparisons of student gains over time.)

Florida's Achievement Gaps on the NAEP in the Last 10 Years

Florida vs. Nation: Achievement Gaps

Over time, U.S. to Florida comparisons show that gaps have closed or are decreasing on NAEP mathematics and reading tests taken by elementary/middle schoolers, with Florida students maintaining an advantage in writing (Table 3, [Appendix F](#)). Table 3 depicts how Florida students changed on NAEP performance between 1992-2003. Changes of 3 scale score points or more have been used to identify increasing or decreasing trends. In writing there was no gap documented when the nation's and Florida's students were compared in 2002: 85% of 4th graders and 84% of 8th graders in both the U.S. and Florida were at or above the Basic level.

Table 3: NAEP Scale Score Gaps — U.S. vs. Florida Over Time

Grade	Subject	Year	U.S. Mean Scale Score	Florida Mean Scale Score	Gap
Grade 4	Math	1992	219	214	+5.0
		2003	234	234	+0.0 (No Gap)
	Reading	1992	215	208	+7.0
		2003	218	216	+2.0 (Decrease)
	Writing	2002	153	158	-5.0 (No Gap*)
Grade	Subject	Year	U.S. Mean Scale Score	Florida Mean Scale Score	Gap
Grade 8	Math	1992	267	260	+7.0
		2003	276	271	+5.0 (Same)
	Reading	1998	261	255	+6.0
		2003	261	257	+4.0 (Same)
	Writing	1998	153	158	-5.0
		2002	152	154	-2.0 (No Gap*)

Florida with advantage. Gap= (U.S. Mean-Florida Mean)

Data Source: NAEP State Profiles, <http://nces.ed.gov/nationsreportcard/states/>

NAEP Gender and Ethnic Gap Trends in Florida.

Within sub-samples of Florida's test-takers over time, a large male to female gap in reading starts in elementary school and widens in middle school; likewise, a male to female gap in mathematics is smaller in elementary school, but also widens in middle school (Table 4). Males performed slightly better in mathematics, while females showed a clear advantage in reading, with the reading gap widening in middle school. Boys tend to post a small advantage in mathematics over girls in both grades 4 and 8; the largest gap is in grade 8 where males were 3.9 points higher on the average scale score. In reading

the trend is reversed: girls did significantly better than boys in both grades, with average scale scores that are 6-13 points higher in grades 4 and 8.

Within sub-samples of Florida's test-takers, gaps remained the same or decreased in NAEP mathematics in White-minority group comparisons over time. Ethnic gaps were found to remain the same or increase on NAEP Reading and Writing tests over time (Table 4). Between 1992 and 2003, the White-Hispanic gap has decreased by 3 or more scale score points in mathematics in grades 4 and 8. The largest decrease in the White-Black gap (of 6 scale score points) was documented between 1992 to 2003, a change from 34 to 28 points. In reading, however, the White-Hispanic gap in grade 4 increased from 14.6 points in 1992 to 18.3 in 2003, an increase in 3 scale score points. It remained the same (below a 3 point difference) in the White-Black comparisons in grades 4 and 8. In writing, the achievement gap could be examined only in grade 8 from 1998 to 2002, given the data available. The White-Hispanic gap increased from 13.7 points in 1998 to 19.0 in 2002.

Table 4: Florida Students in Grades 4 and 8 – Changes in Gender and Ethnic NAEP Gaps in Scale Scores

Grade	Subject	Year	Male-Female	White-Black	White-Hispanic
Grade 4	Math	1992	+2.6	+34.4	+16.0
		2003	+1.9 (Same)	+28.1 (Decrease)	+11.0 (Decrease)
	Reading	1992	-6.1*	+32.9	+14.6
		2003	-8.5* (Increase)	+31.3 (Same)	+18.3 (Increase)
	Writing	2002	-19.3*	+20.4	+10.9
Grade 8	Math	1992	+0.5	+36.3	+26.0
		2003	+3.9 (Increase)	+37.0 (Same)	+17.3 (Same)
	Reading	1998	-13.5*	+28.1	+16.8
		2003	-12.5* (Same)	+29.2 (Same)	+17.3 (Same)
	Writing	1998	-21.7*	+23.8	+13.7
		2002	-25.1* (Same)	+25.5 (Same)	+19.0 (Same)

Data Source: NAEP State Profiles, <http://nces.ed.gov/nationsreportcard/states/>

High School Graduation Rates: U.S. vs. Florida

There is a large gap in high school graduates in U.S.-Florida comparisons. Data show generally declining high school graduation rates in both, with Florida at 55% in 2000, trailing the nation by 10-12% through 2000 (Table 5). Data from 1992-2000, extracted from the NCES Common Core of Data surveys, shows that the proportion of high school students who graduate has gradually declined, both in Florida and in the nation. In Florida, the proportional declines are 65% in 1992, to 57.6% in 1997, reaching 55.2% in 2000. The number of diplomas issued by year, however, has increased over time.

Graduation rates are calculated based on the ratio of the number of regular public school diplomas issued to the number of public school ninth graders enrolled three years before, unadjusted for grade retention and migration for the U.S. as a whole and for each state. The numbers of regular diplomas issued by public schools in the spring for the particular years are shown in parentheses below the graduation rates in Table 5.

**Table 5: Graduation Rates in Florida and the Nation
(Number in Diplomas in Parentheses)**

	1992	1997	2000
Florida	65.0%	57.6%	55.2%
	(93,947)	(95,528)	(107,273)
Nation	71.7%	68.9%	67.2%
	(2,227,562)	(2,403,774)	(2,555,836)

Data Source: U.S. Department of Education, National Center for Educational Statistics, Common Core of Data Surveys (<http://nces.ed.gov>).

Florida's High School Dropout Rates since A+

Florida's dropout rate, published by the Florida Department of Education, shows a slowly declining trend since the A+ legislation was passed (Table 6). Florida's dropout rates spiked at 5.4% in 1999, but have since fallen to around 3% in 2003. On the Website showing school indicators used in Florida, dropout rate is defined as the percentage calculated by dividing (a) the number of students in grades 9-12 who provided withdrawal reasons by (b) the year's total enrollment for grades 9-12. It is assumed that all those who provided withdrawal reasons are the only students who actually withdrew. The dropout calculations also do not account for changes in student numbers due to other factors, such as student retention in grade or migration due to school vouchers (an anticipated consequence of the legislation).

Table 6: Dropout Rates Reported by Florida DOE

Year	Percent Dropouts
1998	3.9%
1999	5.4%
2000	4.6%
2001	3.8%
2002	3.2%
2003	3.1%

Data Source: Florida Department of Education,
Data and Reports <http://info.doe.state.fl.us/>

College-bound Seniors in the Nation and Florida:

SAT Verbal and Math Scores 1992-2003

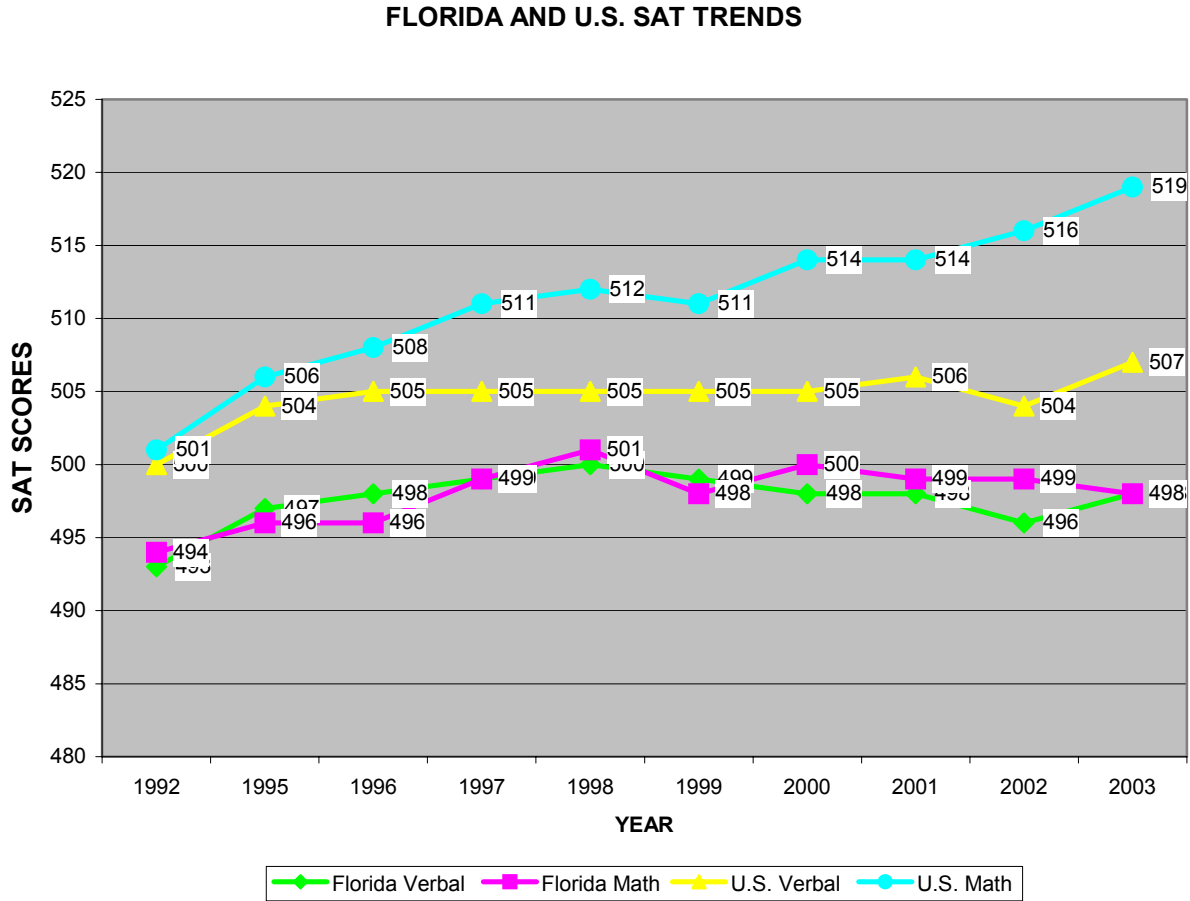
SAT performance should be evaluated in conjunction with information on students who opt to take the test, as there is a documented relationship between student background factors, such as socioeconomic status (SES) and gender, and SAT scores. Based on self-reported family incomes, ethnic minorities in Florida and in the nation tend to come from lower income brackets.⁴⁵

More College-bound seniors from Florida are taking the SAT, with greater percentages from ethnic minority groups ([Appendix G](#)). The number of Florida students taking the SAT increased from 48,570 in 1992 to 83,397 in 2003. Over this period, the number of female students taking the SAT increased by about 1%, the number of Black students increased by about 3%, and the number of Hispanic students by about 5%, with the number of high-income students reduced by 1% proportionately. It should be noted that in 2003, there is a marked gap in terms of SES in SAT test-takers from Florida versus those in the nation, with 58% of the nation's test-takers indicating that they were from high-income families (earning \$ 50,000 or more annually) relative to only 25% in Florida.

U.S. to Florida trends in College Bound Seniors shows a widening gap in both quantitative and verbal areas. Because only 25% of Florida's test-takers report membership in high-income families relative to 58% in the nation, and SAT performance is known to be related to family income, these trends are expected (Figure 4, [Appendix G](#)). Mean SAT Verbal scores in Florida have slowly increased from 493 in 1992 to 500 in 1998, settling at 498 in 2003. Mean SAT mathematics scores show a similar trend and numbers, starting with 494 in 1992, rising to 501 in 1998, and dropping a few points to 498 in 2003. Compared to the nation, the Florida gaps by year have been marginal on the verbal test (less than 1/10 of a standard deviation unit), ranging from -7 points to -9 points. In mathematics, however, the gap has been larger, with -11 points (about 1/10 of a standard deviation unit) in 1998, and increasing to -21 (about 2/10 of a standard deviation unit) in 2003.

Ethnic gaps in SAT scores are wide among Floridians ([Appendix H](#)). In Florida, White-Black gaps in the verbal and mathematics areas have stayed at about 90 points (close to one standard deviation unit below Whites), and the White-Hispanic gap is about 40 points (close to half a standard deviation unit below Whites). The gaps seem to have increased over time by about 3-5 points for both paired groups.

Figure 4: Performance Gaps in U.S. vs. Florida College-bound Seniors



U.S.-Florida Gaps by Ethnic Category on SAT Composite show that Florida's Hispanics have an advantage over Hispanics in the nation despite their generally lower income backgrounds, but other ethnic groups trail the nation (Table 7). In 2003, when comparisons were made between the nation and Florida by ethnic group on the SAT V-M composite, Florida's Hispanic students had a +39 point advantage over Hispanic test-takers in the nation (about 4/10 of a standard deviation unit). This stands in stark contrast to the gap found in all other ethnic group comparisons, which shows Florida students at a disadvantage of 10 to 34 points behind their national peers by ethnic group. More research using individual student data is necessary to verify these general trends.

**Table 7: 2003 SAT Composite (Verbal plus Math Score)
Florida to U.S. Gaps by Ethnic Group***

Ethnicity	Florida	U.S.	Gap
White	1041	1063	-22
African American	847	857	-10
Hispanic	951	912	+39
Asian	1049	1083	-34
Other	975	1006	-31

*Based on test-takers' indicated ethnicity

Data Source: Miller, M. (2003). SAT Trends: Florida and the Nation, 19.
Tallahassee, FL: Department of Education.

Section 4: Conclusions

Good News in Elementary and Middle School: Strongest Performance in Writing

Based on performance on both the FCAT-CRT and NRT, Florida's elementary and middle school students have shown steady gains in reading, mathematics, and writing between 1999 and 2003, while Florida's high schoolers made small but steady increases in mathematics and writing. The results on both state tests concur for all the grades tested, with one test comparing student performance to national norms. Furthermore, the NAEP findings on the same subjects for grades 4 and 8 also agree, lending further support to the conclusion that the observed achievement gains are real. NAEP proficiency levels are set at difficult levels, and the results are particularly encouraging in writing, where the scale score mean for Florida is actually higher than the nation's mean in some cases.⁴⁶ The strongest and most sustained performance in Florida's students by grade and over the years is in the area of writing.

Bad News: High School Reading Achievement Trends

The bad news appears to be at the high school level in reading. More than 60-70% students are below the passing level on the FCAT-CRT. They are also, on average,

falling below the national median in the last few years on the FCAT-NRT. Again, the findings on both state test indicators concur: reading is a weak area for Florida's high school students.

Mixed News: Graduation/Dropout Rates and College-bound Seniors

The U.S.-Florida high school graduation gap also deserves attention. Compared to graduation rates in the nation over time, Florida's graduation rates have stayed *well below* the national rate. Overall, on surveys in both the nation and Florida, graduation rates appeared to be falling through 2000, and historically, fewer students graduate from high school in Florida compared to the nation.

The reported dropout rates seemed to be settling at around 3% in Florida, lower than before. Calculations of both dropout and graduation rates, however, would be more informative if they take into account the new retention policy for students, and migration from school to school due to Florida's new voucher program ("opportunity scholarships") as well as other student instability factors. Studies should focus on earlier grades as well. How many children are being held back in grades 6-8, and if so, how many of those are dropping out or moving before they reach high school? The A+ legislation and concomitant policies may force some of these shifts.

At the same time, more ethnically diverse college-bound seniors are taking the SAT in Florida. In general, Florida students are below the national mean in both verbal and math areas, but more so in mathematics. Because there is a documented relationship between the SAT and the socio-economic status of students, and only 25% of Florida's students reported family incomes of \$50,000 or greater compared to 58% of the nation's test-takers in 2003, this trend is predictable. Hispanics in Florida are doing better on the SAT composite than Hispanics in the nation, however, despite the generally lower

income levels documented in this ethnic group. This is particularly noteworthy and warrants a systematic study of why this is happening.

Mixed to Bad News on Achievement Gaps in Various Subgroups in Florida

Although they remain the same or decrease in some cases, ethnic achievement gaps continue to be large at all levels in Florida on state tests. Similarly, on the NAEP, there are male-female gaps in reading and mathematics (in different directions, with boys doing better in math, girls in reading), gaps that seem to widen as students get older. Compared to Whites, other subgroups such as Limited English Proficient (LEP) and Exceptional Students seem to be far behind on state tests. The gaps seem to be closing in early grades for some groups, but that pattern is not sustained in higher grades. One factor to keep in mind is that the subgroups are not mutually exclusive. Just as a White child can be poor and have exceptional needs, a Hispanic child can be poor and be LEP. This complicates challenges faced by individual children, and schools have to align resources and practices in different ways to meet their needs. The goal of closing the achievement gap in all subgroups within inflexible time frames may not be realistic or achievable.

Section 5: Recommendations

Florida's reforms have led to moderate improvements in the last 10 years. The trend data show positive achievement changes beginning in elementary levels in Florida schools since the A+ program was implemented; but the data also show that these changes are not sustained at the secondary level. What can be done to continue progress along the same path, with sustained growth and reductions in observed achievement gaps at all levels, and in an environment where schools and teachers feel more supported?

Given the current data trends and the general principles regarding best practices for standards-based educational reforms discussed earlier in this brief, it is recommended that Florida legislators and policy-makers undertake the following:

1. Provide local support and evaluation training opportunities to help teachers and school leaders make better use of data to reduce achievement gaps in reading and mathematics, particularly at the secondary level.
2. Document and reward the pedagogically sound practices that follow from Recommendation 1, even when short-term gains are small. These practices, informed by data, will support student learning and sustained growth.
3. Create conditions to help schools teach to the Sunshine State Standards along a K-12 continuum -- conditions that lead to real engagement of teachers and students in learning activities tied to standards and that reduce the amount of test preparation. This includes removing the use of teaching to the (FCAT) test language, currently on the A+ website, that tends to encourage short-term gains.

4. Make greater use of appropriate sampling techniques (NAEP type of design) in school evaluations, lowering anxiety and fear resulting from the present accountability system. This will allow Florida to retain high levels of accountability while lowering the stakes for individual school staff.
5. Seek ways to set reasonable and staggered timelines, within the NCLB and A+ frameworks, for schools with diverse and challenging populations to demonstrate adequate yearly progress (AYP).
6. Rethink sanctions and grade retention policies in light of new and past research that finds that grade retention does not improve student achievement.
7. Identify specific reforms with positive and sustained achievement effects by supporting long-term research that looks at diverse students' achievement in relation to school and teacher practices over time.

Notes and References

- ¹ Buttram, J. & Waters, J. (1997). Improving America's Schools Through Standards-based Education. *National Association of Secondary School Principals*, 81(590), 1-6.
- ² State of Florida website. Retrieved January 16, 2004, from <http://www.myflorida.com/myflorida/education/learn/aplusplan/youKnow.html>
- ³ See, for example:
- Cohen, D. (1995). What Is the System in Systemic Reform? *Educational Researcher*, 24(9), 11-17.
- Cohen, D. & Hill, H. C. (1998). State Policy and Classroom Performance: Mathematics Reform in California. Philadelphia: Consortium for Policy Research in Education, Graduate School of Education, University of Pennsylvania.
- ⁴ State of Florida website. Retrieved January 16, 2004, from <http://www.myflorida.com/myflorida/education/learn/aplusplan/youKnow.html>
- ⁵ See, for example:
- Abrams, L. M. (2004, February). The Effects of High-stakes Testing on Teaching and Learning: What Florida Teachers Think. *Education Policy Research Unit*, Doc No. EPSL-0401-104-EPRU. Tempe, AZ: Education Policy Studies Laboratory.
- Shepard, L. A. (1990). Inflated Test Score Gains: Is the Problem Old Norms or Teaching to the Test? *Educational Measurement: Issues and Practice*, 9(3), 15-22.
- Shepard, L. & Dougherty, K. (1991, April). *Effects of High-stakes Testing on Instruction*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- ⁶ State of Florida website. Retrieved January 16, 2004 from <http://www.myflorida.com/myflorida/education/learn/aplusplan/youKnow.html>
- ⁷ See, for example:
- Sanders, W. L., & Horn, S. P. (1994). The Tennessee Value-Added Assessment System (TVAAS): Mixed-model Methodology in Educational Assessment. *Journal of Personnel Evaluation in Education*, 8(2), 299-311.
- Sanders, W. & Horn, S. (1998). Research Findings from the Tennessee Value-Added Assessment System Database. *Journal of Personnel Evaluation in Education*, 12(3), 247-256.
- State of Florida website. Retrieved January 16, 2004, from <http://www.myflorida.com/myflorida/education/learn/aplusplan/youKnow.html>
- ⁸ No Child Left Behind Act of 2001, Public Law 107-110, signed into law January 8, 2002: 115 Stat. 1425.
- ⁹ Schemo, D. (2001, February 16). Threat of Vouchers Motivates Schools to Improve, Study Says. *New York Times*.
- ¹⁰ National Commission of Excellence in Education (1983). A Nation at Risk: The Imperative for Educational Reform. Washington, DC: U.S. Government Printing Office.
- ¹¹ No Child Left Behind Act of 2001, Public Law 107-110, signed into law January 8, 2002: 115 Stat. 1425.

¹² See, for example:

Stecher, B. & Barron, S. (1999). *Test-based Accountability: The Perverse Consequences of Milepost Testing*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada.

Stecher, B., Barron, S., Chun, T., & Ross, K. (1999). *The Effects of Washington State Education Reform on Schools and Classrooms* (CSE Technical Report No. 525). Los Angeles: National Center for Research on Evaluation, Standards and Student Testing (CRESST), UCLA.

Stecher, B., Barron, S., Kaganoff, T., & Goodwin, J. (1998). *The Effects of Standards-based Assessment on Classroom Practices: Results of the 1996-1997 RAND Survey of Kentucky Teachers of Mathematics and Writing* (CSE Technical Report No. 482). Los Angeles: Center for Research on Evaluation, Standards and Student Testing (CRESST), UCLA.

¹³ No Child Left Behind Act of 2001, Public Law 107-110, signed into law January 8, 2002: 115 Stat. 1425.

¹⁴ Gandal, M. (1995). Why We Need “Academic” Standards. *Educational Leadership*, 53(1), 84-86.

¹⁵ Gandal, M. & Vranek, J. (2001). Standards: Here Today, Here Tomorrow. *Educational Leadership*, 59(1), 6-13.

¹⁶ *Ibid.*

¹⁷ Linn, R., Baker, E., & Betebenner, D. (2002). Accountability Systems: Implications of Requirements of the No Child Left Behind Act of 2001. *Educational Researcher*, 31(6), 3-16.

¹⁸ *Ibid.*

¹⁹ Linn, R. L. (2003). Accountability: Responsibility and Reasonable Expectations. *Educational Researcher*, 32(7), 3-13.

²⁰ *Ibid.*

²¹ See, for example:

Abrams, L. M. (2004, February). The Effects of High-stakes Testing on Teaching and Learning: What Florida Teachers Think. *Education Policy Research Unit*, Doc No. EPSL-0401-104-EPRU. Tempe, AZ: Education Policy Studies Laboratory

Mehrens, W., & Kaminski, J. (1989). Methods for Improving Standardized Test Scores: Fruitful, Fruitless, or Fraudulent? *Educational Measurement: Issues and Practice*, 8(1), 14-22.

Shepard, L. A. (1990). Inflated Test Score Gains: Is the Problem Old norms or Teaching to the Test? *Educational Measurement: Issues and Practice*, 9(3), 15-22.

Hamilton, L., Stetcher, B., & Klein, D., (Eds.) (2002). *Making Sense of test-based Accountability in Education*. Washington DC: RAND.

Taylor, G., Shepard, L., Kinner, F., & Rosenthal, J. (2003). *A Survey of Teachers' Perspectives on High-stakes Testing in Colorado: What Gets Taught, and What Gets Lost* (CSE Technical Report 588). Los Angeles: Center for Research on Evaluation, Standards and Student Testing (CRESST), UCLA.

²² Pedulla, J., Abrams, L., Madaus, G., Russell, M., Ramos, M., & Miao, J. (2003). *Perceived Effects of State-mandated Testing Programs on Teaching and Learning: Findings from a National Survey of*

- Teachers*. Chestnut Hill, MA: National Board on Educational Testing and Public Policy, Boston College.
- ²³ George, P. (2001). A+ Accountability in Florida? *Educational Leadership*, 59(1), 28-32.
- ²⁴ Cizek, G. (2001). More Unintended Consequences of High-stakes Testing. *Educational Measurement: Issues and Practice*, 20(4), 19-27.
- ²⁵ Cizek, G. (2001). More Unintended Consequences of High-stakes Testing. *Educational Measurement: Issues and Practice*, 20(4), 19-27.
- ²⁶ *Ibid.*
- ²⁷ American Federation of Teachers (AFT) (2001). Making Standards Matter, 2001. *American Educator*, 25(4), 47-48.
- ²⁸ Doboer, G. (2002). Student-centered Teaching in a Standards-based World: Finding a Sensible Balance. *Science and Education*, 11, 405-417
- Harrington-Lueker, D. (1998). States Raise the Bar: Now Local School Districts Are Accountable for Results. *American School Board Journal*, 185(6), 17-21.
- Kluth, P. & Straud D. (2001). Standards for Diverse Learners. *Educational Leadership*, 59(1), 43-46.
- ²⁹ Buttram, J. L., & Waters, J. T. (1997). Improving America's schools Through Standards-based Education. *National Association of Secondary School Principals*, 81(590), 1-6.
- ³⁰ American Federation of Teachers (AFT) (2001). Making Standards Matter, 2001. *American Educator*, 25(4), 47-48.
- ³¹ Glass, G. V (2002). Teacher Characteristics. In A. Molnar (Ed.). *School Reform Proposals: The Research Evidence*. Research in Educational Productivity Series (Ch. 8). Greenwich, CT: Information Age Publishing, Inc.
- ³² Walberg, H. J. (2003). Policy Brief 3: Closing the Achievement Gap. *Pre-publication version*. Jacksonville, FL: University of North Florida.
- ³³ Harrington-Lueker, D. (1998). States Raise the Bar: Now Local School Districts are Accountable for Results. *American School Board Journal*, 185(6), 17-21.
- ³⁴ Viadero, D. (2004). Achievement-gap study Emphasizes Better Use of Data. *Education Week*, 23(19), 9.
- ³⁵ Chatterji, M. (2002). Models and Methods for Examining Standards-based Reforms and Accountability Initiatives: Have the Tools of Inquiry Answered Pressing Questions on Improving Schools? *Review of Educational Research*, 72, 345-386. Recommends systemic models of evaluation for improving schools and school programs.
- ³⁶ Deschenes, S., Cuban, L., & Tyack, D. (2001). Mismatch: Historical Perspectives on Schools and Students Who Don't Fit Them. *Teachers College Record*, 103, 525-547.
- ³⁷ Morris, D. (2001). Assessing the Implementation of High-stakes Reform: Aggregate Relationships between Retention Rates and Test Results. *National Association of Secondary School Principals*, 85(629), 18-34.
- ³⁸ *Ibid.*
- ³⁹ Walberg, H. (2003). Policy Brief 3: Closing the Achievement Gap. *Prepublication version*. Jacksonville, FL: University of North Florida.

⁴⁰ Loveless, T. (1999). *The Tracking Wars*. Washington DC: The Brookings Institution Press.

Campbell, J., Hombo, C., & Masseo, J. (2000). *NAEP 1999 Trends in Academic Progress: Three Decades of Student Performance*. Washington DC: Office of Educational Research and Improvement (OERI), U.S. Department of Education.

⁴¹ Lee, J. (2002). Racial and Ethnic Achievement Gap Trends: Reversing the Progress Toward Equity? *Educational Researcher*, 31(1), 3-12.

⁴² *Ibid.*

⁴³ *Ibid.*

⁴⁴ *Ibid.*

⁴⁵ Miller, M. (2003). *SAT trends: Florida and the Nation*. Tallahassee, Florida: Florida Department of Education.

⁴⁶ Linn, R., Baker, E., & Betebenner, D. (2002). Accountability Systems: Implications of Requirements of the No Child Left Behind Act of 2001. *Educational Researcher*, 31(6), 3-16.